The Anti-tumor Toxin CD437 is a Direct Inhibitor of DNA Polymerase α

Ting Han[1], **Maria Goralski**[2], **Emanuela Capota**[2], **Shae B. Padrick**[3], **Jiwoong Kim**[4,5], **Yang Xie**[4,5], and **Deepak Nijhawan**[1,2,5,*]

[1]Department of Biochemistry, UT Southwestern Medical Centers, 5323 Harry Hines Boulevard, Dallas, TX 75390-9152

[2]Department of Internal Medicine, UT Southwestern Medical Centers, 5323 Harry Hines Boulevard, Dallas, TX 75390-9152

[3]Department of Biophysics, UT Southwestern Medical Centers, 5323 Harry Hines Boulevard, Dallas, TX 75390-9152

[4]Quantitative Biomedical Research Center, Department of Clinical Sciences, UT Southwestern Medical Centers, 5323 Harry Hines Boulevard, Dallas, TX 75390-9152

[5]Harold C. Simmons Comprehensive Cancer Center, UT Southwestern Medical Centers, 5323 Harry Hines Boulevard, Dallas, TX 75390-9152

Abstract

CD437 is a retinoid-like small molecule that selectively induces apoptosis in cancer but not normal cells through an unknown mechanism. We used a forward genetic strategy to discover mutations in *POLA1* that coincide with CD437 resistance (*POLA1^R*). Introduction of one of these mutations into cancer cells by CRISPR/Cas9 genome editing conferred CD437 resistance demonstrating causality. *POLA1* encodes DNA polymerase α, the enzyme responsible for initiating DNA synthesis during the S phase of the cell cycle. CD437 inhibits DNA replication in cells and recombinant POLA1 activity *in vitro*. Both effects are abrogated by mutations associated with *POLA1^R*. In addition, we detected an increase in the total fluorescence intensity and anisotropy of CD437 in the presence of increasing concentrations of POLA1 consistent with a direct binding interaction. The discovery of POLA1 as the direct anti-cancer target for CD437 has the potential to catalyze its development into an anti-cancer therapeutic.

---

*To whom correspondence should be addressed: Deepak Nijhawan, Telephone: 214-648-6713, deepak.nijhawan@utsouthwestern.edu.

Introduction

CD437, a retinoid-like small molecule, is a promising cancer drug lead based on its potential to achieve a high therapeutic index (Figure 1A). CD437 is toxic to numerous cancer cell lines, which have been derived from different primary tumors including ovarian cancer, non-small cell lung cancer, leukemia, breast cancer, and squamous cell carcinoma[1-5]. After CD437 treatment, cancer cells accumulate in the S-phase of the cell cycle and then undergo apoptosis characterized by loss of mitochondrial membrane potential, cytochrome c release, and caspase activation[6,7]. By contrast, CD437 does not induce apoptosis in normal human cells. Rather, both primary keratinocytes and normal human lung epithelial cells undergo arrest in the G1/S phase of the cell cycle, which is not followed by cell death[8,9]. Indeed, peritoneal administration of CD437 over three weeks led to regression of mouse xenograft tumors derived from human cancer cell lines without any signs of toxicity[3]. The promise of CD437 as a therapeutic lead inspired extensive chemical optimization yielding both potent and bioavailable derivatives, some of which are efficacious following oral delivery[10,11]. Further development of CD437, however, has been stymied by a lack of knowledge of either its direct protein target or its mode of action.

Although CD437 was first identified as a selective agonist of the gamma isoform of retinoid acid receptor (RARγ)[12], several lines of evidence suggest that CD437-induced cell death proceeds through an alternative target. First, non-small cell lung cancer and breast cancer cell lines that are differentially sensitive to endogenous RAR agonists, such as retinoic acid, are universally sensitive to CD437[4]. Second, co-incubation of CD437 with an antagonist to all three RAR isoforms (RARα, RARβ, RARγ) did not inhibit toxicity[13,14]. Third, through structure-activity relationships of CD437, no correlation was observed between the ability of an analog to activate RARγ and cytotoxicity[10].

If the anti-tumor activity of CD437 depended on the activation of RARγ, the expectation would be that cells lacking RARγ expression would be insensitive to CD437. In fact, the converse is true: cancer cells that do not express RARγ are at least equally and in some cases more sensitive to CD437. For instance, leukemic cells that express no functional RARs remained sensitive to CD437[15]. The importance of RARγ engagement in CD437 toxicity was also directly tested in F9 teratocarcinoma cells, which are sensitive to CD437 and express RARγ. Through homologous recombination several F9 clones were isolated with homozygous loss of the RARγ gene[16]. In spite of no RARγ expression, these knockout cells were equally sensitive to CD437 and related analogues *in vitro*. *In vivo*, tumors derived from RARγ -/- cells were, in fact, more responsive to CD437 treatment. A paradoxical relationship between RARγ expression and CD437 was also found in an analysis of gene expression and CD437 sensitivity in more than 800 human cancer cell lines representing multiple lineages[5]. In this large dataset, there was a significant correlation between cells with lower RARγ expression and CD437 sensitivity. Notwithstanding, none of the aforementioned observations and experiments offers any evidence in support of the claim that CD437 is toxic via its RARγ agonist activity. To date, the mode of action for the anti-tumor activity of CD437 remains unknown.

Here, we have used a genetic system to identify compound-resistant alleles of *POLA1*, which engender CD437 resistance. *POLA1* encodes the catalytic subunit of DNA polymerase α, which is required for the initiation of DNA replication. Using a combination of biochemistry and biophysics, we provide evidence that CD437 exerts its cytotoxicity by directly binding and inhibiting POLA1.

Results

POLA1 mutations coincide with CD437 resistance

We used HCT-116 cells to identify such compound resistant alleles based on a recently described forward genetics approach[17]. HCT-116 is a colorectal cancer cell line that lacks expression of MLH1, which is a protein essential for DNA mismatch repair. These cells have a high nucleotide substitution rate[18], which serves as a mechanism of mutagenesis and predisposes the cells to develop resistance to toxins as a consequence of heterozygous mutations.

A population of HCT-116 cells is expected to contain multiple resistant "founders." Each founder represents a mutational event and gives rise to a family of resistant clones (Supplementary Results, Supplementary Figure 1A). To distinguish between founder events, we selected for CD437 resistance amongst a barcoded population of HCT-116 cells. To barcode the cells, we engineered a lentiviral vector library to contain ~1000 different oligonucleotides and used it to package lentivirus and infect HCT-116 cells (Supplementary Figure 1B). We subjected 10 million barcoded HCT-116 cells to selection with CD437 and isolated 20 resistant clones. For each clone, we deciphered the barcode by Sanger sequencing a PCR product, which was amplified from genomic DNA using primers that encompass the aforementioned oligonucleotide. Using these sequences, we were able to cluster the 20 toxin-resistant clones into 10 CD437-resistant families (Supplementary Figure 1C). Since the diversity of our original plasmid library is $\sim 10^3$, we predicted that each of these families represented an independent mutational event.

We analyzed one member from six independent CD437-resistant families for toxin resistance. These clones were between 3 and 6-fold less sensitive to the toxin than the parental cell line (Figure 1B, Supplementary Figure 1E). We counter-screened each of these clones for resistance to an unrelated toxin, paclitaxel, which is a substrate for multiple drug efflux pumps. None of the clones were resistant to paclitaxel, reducing the likelihood that CD437 resistance could be explained simply by non-specific toxin metabolism or efflux (Supplementary Figure 1D, 1E).

We hypothesized that resistance in these clones might be the result of compound resistant alleles in the CD437 target. In order to identify these mutations, we subjected the six CD437-resistant clones and 13 CD437-sensitive clones to whole exome-sequencing at a mean depth between 84× and 187× coverage (Supplementary Table 1, Supplementary Dataset 1). In our analysis of sequencing results, we made the assumption that CD437 resistant alleles were less likely to result from nonsense mutations or indels (insertions/ deletions). Therefore, we restricted our analysis to non-synonymous mutations that were present in the six CD437 resistant clones, but not in the 13 CD437 sensitive clones

(Supplementary Dataset 2). Using this approach, we found 772 genes that had a missense mutation in at least one of the six CD437 resistant clones (Figure 1C). Since the CD437 target was likely mutated in different founder clones, we counted the numbers of genes that were recurrently mutated within the group of six. A single gene, *POLA1*, harbored mutations in all six clones and therefore was the leading candidate for the target of CD437 (Figure 1C). *POLA1* is located on the X chromosome and is monoallelic in the male derived HCT-116 cells.

The six identified mutations in *POLA1* affect five amino acids clustered between position 691 and 772 (C691Y, L700S, L764S, I768T, A772D, and A772T) outside of the catalytic domain (Figure 1D). Each of the six mutations led to a non-conservative substitution. The cysteine at position 691 was mutated to tyrosine, and the other five mutations all resulted in the substitution of a hydrophobic for a hydrophilic amino acid. We mapped each of these mutations to the recently solved X-ray crystal structure of human POLA1[19]. L764, I768, and A772 all map to the same surface of an alpha helix, and are positioned near the other two mutated residues, L700 and C691 (respectively, L700: 3.9, 6.1, 14.1 Å and C691: 10, 12, 17 Å) (Figure 1E). The clustering of these residues in three-dimensional space supports the hypothesis that these mutations are indeed related to CD437 resistance. In particular, we reasoned that the mutation of alanine 772 was unlikely to happen by chance, as it was discovered in two independent resistant clones resulting in two different substitutions (A772T and A772D).

POLA1 mutations cause CD437 resistance

To confirm that mutations in *POLA1* do indeed cause CD437 resistance, we introduced the L764S allele into HCT-116 cells using CRISPR/Cas9 technology. We chose this mutation because the codon is positioned adjacent to a consensus protospacer adjacent motif (PAM) sequence, which is required for CRISPR genome editing (Figure 2A). We transfected HCT-116 cells with either GFP (mock), hCas9 and a guide RNA targeting *POLA1* at L764 (sgRNA), or the sgRNA, hCas9 and a single stranded oligodeoxynucleotide (ssODN). The ssODN flanks the predicted double strand break site and is designed to serve as an alternative repair template, encoding the L764S allele. In comparison to either GFP-transfected cells or cells transfected with hCas9 and sgRNA, the addition of the ssODN transfection led to a higher rate of CD437 resistance (Figure 2B). We isolated four independent clones and analyzed the sequence of *POLA1* surrounding position 764 by Sanger sequencing. As expected all four clones harbored the L764S allele encoded by the ssODN template (Supplementary Figure 2A). In addition, all four clones exhibited between 26 and 43-fold resistance when tested for sensitivity across a range of CD437 concentrations (Figure 2C). The high nucleotide substitution rate in HCT-116 cells introduced the possibility that the selected clones harbored additional mutations that contribute to resistance.

To address this possibility, we repeated the experiment in HeLa cells, which are not known to have defects in mismatch repair or a high nucleotide substitution rate. HeLa cells transfected with Cas9 and sgRNA had a higher rate of resistance than the GFP transfected cells, and the rate was further increased when the ssODN was included in the transfection

(Figure 2D). We isolated four clones from the lattermost condition and analyzed the *POLA1* sequence. HeLa cell clones 1 and 3 harbored mutations encoded by the ssODN (L764S), however, clones 2 and 4 harbored a L762R and L764F mutation, which were not coded by the ssODN (Supplementary Figure 2B). These mutations are possibly the result of imprecise template-independent Non Homologous End Joining (NHEJ) of the double stranded DNA break generated by hCas9. This is consistent with the observed CD437 resistance in cells transfected with hCas9 and sgRNA without the ssODN template. Regardless, all four HeLa cell clones were between 14 and 29- fold resistant to CD437 in comparison to the parental HeLa cell population (Figure 2E). The identification of L762R and L764F serves to reinforce the relationship between *POLA1* mutations and CD437 resistance.

CD437 is cytotoxic by inhibiting POLA1—*POLA1* encodes DNA polymerase α, an enzyme that initiates DNA synthesis by using an RNA primer to synthesize the first ∼10-20 base pairs of DNA, thereby providing the substrate for more processive DNA polymerases. To functionally test whether CD437 influences DNA polymerase α activity, we measured bromodeoxyuridine (BrdU) incorporation after a two hour CD437 treatment in either the parental or CD437-resistant HCT-116 cell lines. In the parental cell line, we found that 1.5 μM CD437 blocked BrdU incorporation. By contrast, resistant cell lines harboring any one of the five different mutations required at least 15 μM of CD437 to achieve the same level of inhibition (Figure 3A, 3B).

Additionally, we tested the effect of CD437 on DNA polymerase α activity *in vitro* using recombinant POLA1 proteins (Supplementary Figure 3A) and several synthetic nucleic acid substrates (Supplementary Figure 3B). To generate the synthetic substrates for POLA1, we annealed a fluorescently labeled 15 nt RNA to complementary DNA templates of length ranging from 25 nt to 40 nt (Supplementary Figure 3B, 3C). We independently co-incubated each substrate with dNTPs and a recombinant, catalytically active fragment of wild type POLA1[19] and used polyacrylamide gel electrophoresis to analyze the levels of fully extended primer (Supplementary Figure 3D). Wild type POLA1 exhibited robust primer extension activity for all substrates tested (Supplementary Figure 3E, 3F). In addition, co-incubation with CD437 inhibited the activity of POLA1 regardless of template length. The activity of wild type POLA1 and mutant POLA1 in this assay was comparable (Supplementary Figure 3F, 3G). However, wild-type POLA1 was readily inhibited by CD437 (IC_{50} = 22 nM) whereas POLA1 harboring either L700S or the L764S mutation, required a higher concentration of CD437 to inhibit activity (IC_{50} = 1.2 μM and 2.6 μM, respectively) (Figure 3C, 3D, and Supplementary Figure 4). We concluded that cells expressing L700S or L764S POLA1 alleles are resistant to CD437 because these mutations render the enzyme less sensitive to CD437 inhibition. Taken together, we concluded that CD437 is toxic to cells by virtue of its ability to inhibit POLA1.

CD437 toxicity is S-phase dependent—POLA1 functions in DNA replication during the S-phase of the cell cycle, therefore, we tested whether the toxicity of CD437 depends on progression through S-phase. We incubated cells with high concentrations of thymidine to prevent cells from entering or progressing through S-phase (commonly referred to as "thymidine block"). Under these conditions, in comparison to asynchronous cells, the

percentage of cells in the G1-phase of the cell cycle is increased (42% to 78.5%) and those in S-phase are decreased (36% to 22.4%) (Supplementary Figure 5). CD437 treatment of asynchronous HeLa cells resulted in the cleavage of caspase-3, a marker for programmed cell death (Figure 3E). By contrast, in thymidine-blocked cells, CD437 did not undergo cell death evidenced by no caspase-3 cleavage. This observation is consistent with our hypothesis that toxicity of CD437 results from inhibiting POLA1 during S-phase. After releasing cells from thymidine block, cells enter S phase in a coordinated fashion which peaks at four hours (76.5%) (Supplementary Figure 5). We treated cells with CD437 simultaneous with release from thymidine block and, in comparison to asynchronous cells, observed a substantial increase in caspase-3 cleavage (Figure 3E). Taken together, we concluded that CD437 does not prevent cells from entering S-phase, and that the cytotoxicity occurs during the S-phase of the cell cycle when DNA polymerase α is expected to be active.

CD437 directly binds POLA1—CD437 is a fluorescent compound, which allowed us to measure binding by fluorescence spectroscopy[20]. We observed an increase in the fluorescence intensity of CD437 in a POLA1 concentration dependent manner (Figure 4A). Curving fitting to a binding isotherm suggested the dissociation constant (Kd) to be 67 ± 22 nM. Concordantly, we also observed an increase in fluorescence anisotropy of CD437 in the presence of POLA1, indicating that CD437 is bound by the macromolecule POLA1 (Figure 4B). The affinity of CD437 for POLA1 is consistent with the potency of inhibition observed *in vitro* (Figure 3C, 3D). Surprisingly, POLA1 harboring the L764S mutation binds to CD437 with similar affinity as WT POLA1 (94 ± 18 nM), and the L700S mutant binds CD437 with substantially higher affinity (Kd < 5 nM) (Figure 4A). These results suggest that the compound resistant mutations we identified do not block CD437 binding to POLA1. Rather, CD437 may antagonize POLA1 at an allosteric site, and the resistant mutations adopt a confirmation independent of allostery.

Discussion

The most appealing feature of CD437 as a cancer drug lead is that it has selective toxicity toward cancer cells. The inhibition of POLA1 by CD437 leads to a reversible cell cycle arrest in normal cells but induces fulminant apoptosis in cancer cells. One model for this differential response is that POLA1 inhibition by CD437 in normal cells leads to stalling of the replication fork and triggers a signal halting progression either into or through the S-phase of the cell cycle. By contrast, cancer cells may lack this checkpoint and as a consequence progress through S-phase even though DNA replication cannot proceed. In this scenario, the resulting genomic instability may induce an apoptotic signal leading to cell death. There is some evidence suggesting that the CD437 treatment activates the DNA damage response (DDR), which constitutes a signal transduction pathway that induces cell cycle arrest in response to both DNA damage and replication stress[21]. Cells treated with CD437 exhibit increased levels of phosphorylated gamma H2AX, a *bona fide* marker of the DDR[22]. Therefore, one hypothesis is that the stalled replication fork leads to activation of the DDR that triggers checkpoint activation in normal cells but apoptosis in cancer. More

investigation is needed to determine exactly how normal cells sense the inhibition of POLA1 by CD437 and why this signal specifically leads to cancer cell death.

A surprising finding from our study is that POLA1 harboring the L764S mutation binds to CD437 with the same affinity as the wild type protein, even though, the mutant protein is not inhibited by CD437. CD437 is a fluorescent compound, which allowed us to measure binding by analyzing the change in total fluorescence as well as anisotropy in the presence of varying concentrations of POLA1. The changes in total fluorescence were different in the mutant proteins when compared to the wild type. Our hypothesis is that these differences reflect an alternative chemical environment for CD437 bound to mutant POLA1, which may result from either binding to a different site on POLA1 or binding POLA1 in a different confirmation. The latter seems more likely considering the nearly identical dissociation constants between the wild type protein and L764S mutant. Furthermore, these observations raise the hypothesis that CD437 binds POLA1 in a heretofore-unknown allosteric site rather than the active site. A co-crystal structure of CD437 with both mutant and wild type POLA1 will help answer this question.

There are therapeutic implications for CD437 as a newly discovered inhibitor of POLA1 and DNA replication. Many cancer drugs act by targeting any of a variety of aspects of DNA replication. Three classes of agents block DNA replication by influencing the template: anthracyclines through DNA intercalation, topoisomerase inhibitors through alteration of the supercoiled state of DNA, and alkylators through covalent modification of DNA. Other approved drugs inhibit replication by affecting nucleotides, which are essential for DNA synthesis. For instance, hydroxyurea, methotrexate, and pemetrexed reduce the overall level of nucleotides, and gemcitabine and cytarabine block replication by substituting for deoxycytidine. Finally, drugs that target enzymes that regulate DNA replication, such as inhibitors of cyclin-dependent kinases (CDK4 and CDK6), also have therapeutic efficacy[23]. Surprisingly, despite the demonstrable clinical utility of anti-cancer therapies that target essential components of DNA replication, there are no currently available drugs that directly inhibit DNA polymerase.

DNA polymerase α, encoded by *POLA1*, is an ideal enzymatic target for drugs that block DNA replication because it acts at the initial step of DNA synthesis. Aphidicolin is a tetracyclic diterpene antibiotic isolated from the fungus, *Cephalosporum aphidicola*[24], and heretofore was the only described inhibitor of DNA polymerase α. The clinical development of aphidicolin as an anti-cancer agent, however, has been hampered by its poor bioavailability. A phase I trial using aphidicolin glycinate required continuous infusion of the compound in order to achieve the predicted efficacious concentration[25]. Moreover, efforts to synthesize derivatives of aphidicolin that are both active and bioavailable have thus far been unsuccessful[26,27].

Here, we discovered that CD437 is also an inhibitor of DNA polymerase α. Cancer cells have a different response to inhibition of POLA1 by CD437 in comparison to aphidicolin. CD437 mediated inhibition of POLA1 leads to cell death in cancer cells, but induces cell cycle arrest in normal epithelial cells[1-4,6,7,9,28]. In contrast, aphidicolin leads to a reversible cell cycle arrest in both normal and cancerous cells[29]. The unique ability of CD437 to

selectively induce apoptosis in cancer cells may qualify it as a more attractive lead molecule for an anti-cancer therapeutic that directly targets POLA1. CD437 has other advantages as a lead molecule for drug development. Unlike aphidicolin, which is a complex natural product, CD437 is a synthetic molecule that can more readily be optimized via the synthesis of improved analogs. An important milestone in this effort will be the determination of the co-crystal structure of CD437 and POLA1 allowing for structure-aided chemical optimization.

Online Methods

Cell culture

The identity for all human cell lines in this study was confirmed by short tandem repeat (STR) analysis. In addition, all human cell lines were confirmed to be mycoplasma free using a PCR based assay (Genatlantis). HCT-116 cells (obtained from ATCC) were grown in Dulbecco's Modified Eagle's Medium (DMEM) (from Sigma) supplemented with 10% FBS, 2 mM L-Glutamine, and 1× Penicillin-Streptomycin. 293T cells (obtained from Clontech) were cultured in DMEM medium (Sigma) supplemented with 10% FBS. HeLa cells (a gift from Dr. Gelin Wang) were cultured in DMEM media (from Sigma) supplemented with 10% FBS and 2 mM L-Glutamine. These adherent cell lines were cultured employing standard procedures. Sf9 cells (a gift from Dr. Hongtao Yu) were cultured in suspension in Sf-900™ II SFM media (Life Technologies).

Chemicals

CD437 (purified by HPLC to greater than 98% purity) and 5-bromo-2′-deoxyuridine (BrdU) were purchased from Sigma-Aldrich. Paclitaxel was purchased from Selleckchem.

Generation of barcoded HCT-116 cells

60 pmol of 5′-GGAAAGGACGAAACACCGGTT-3′ and 60 pmol of 5′-cgagaattcNNNNNNNNNNNNNNNNNNNNNAAAAAACCGGTGTTTCGTCCTTTCC-3′ were annealed and converted to blunt-ended double-stranded duplexes in a 100 μL reaction using Platinum®Taq DNA Polymerase (Life Technologies). First they were incubated in a thermocycler at 94°C for 5 min, followed by gradual cool down to 55°C (1% ramp), and a further 72°C incubation for 30 min. PCR product was ethanol precipitated, digested with AgeI and EcoRI, and ligated into the pLKO.1 vector (Addgene 10878). Ligated plasmids were transformed into *E. coli* DH5α competent cells and ~1000 colonies resulting from transformation were pooled together to generate a barcoded plasmid library. Lentiviral packaging of the plasmid library was performed by co-transfecting the library with psPAX2 (Addgene 12260) and pMD2.G (Addgene 12259) into 293T cells. Media collected from transfected 293T cells was used to infect 100,000 HCT-116 cells at low multiplicity of infection (MOI) to achieve ~30% cell survival after 3 days of selection with 2 μg/ml puromycin. This resulted in a population of barcoded cells. These cells were expanded for at least 15 population doublings to allow for random mutagenesis.

Selection of resistant clones

For the selection of resistant clones, ten 10 cm plates of barcoded HCT-116 cells (1 million cells per plate) were treated with 2.6 μM CD437. Media containing CD437 was replenished every 3-4 days over the course of 4 weeks. Surviving clones were expanded and a small aliquot of cell pellets for each clone was collected and barcode-genotyped. Based on the barcode sequence, six clones with unique barcodes were subjected to dose response analysis to validate their resistance to the compound used for their selection, as well as to rule out cross-resistance to paclitaxel. All the clones were cryopreserved and the clones validated by dose response analysis were processed for whole exome sequencing.

Barcode amplification and sequencing

Cell pellets (∼0.1 million cells) were resuspended in 50 μL of lysis buffer (10 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.1 mg/ml gelatin, 0.45% NP-40, 0.45% Tween-20, and 0.5 μL of Proteinase K (Thermo Scientific Fermentas FEREO0491), then incubated at 55°C for 2 hours, which was followed by 15 min of incubation at 95°C to inactivate Proteinase K. 2 μL of lysate was used for barcode amplification in a reaction containing 2.5 μL of 10× Standard Taq buffer, 2 μL of 2.5 mM dNTPs, 1 μL of Primer mix (5 μM each of 5′-GGAGGCTTGGTAGGTTTAAGAA-3′ and 5′-GGATCTCTGCTGTCCCTGTAAT-3′), 0.3 μL of Taq polymerase (NEB), and 17.2 μL of H_2O. PCR was performed by 35 cycles of incubation at 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec. PCR product was cleaned up with a DNA clean and concentrator kit (Zymo) and sequenced with a primer 5′-GGAGGCTTGGTAGGTTTAAGAA-3′.

Cell viability assay

4000 HCT-116 cells were plated per well in 96-well assay plates. After overnight incubation, serially diluted compounds were added to the plates to set up a 12-point dose response. A cell survival assay was performed 3 days later using CellTiter-Glo Luminescent Cell Viability Assay kit (Promega), which measures cellular ATP content. The CellTiter-Glo reagent was diluted by adding 1 volume of PBS containing 1% Triton X-100. 40 μL of diluted reagent was directly added to 100 μL of cell culture medium per well. We recorded luminescence using the EnVison multimode plate reader (Perkin Elmer). The IC_{50} was determined from a curve derived from an asymmetric (five parameter) equation and least squares fit in GraphPad Prism.

Next generation sequencing

Genomic DNA purification was performed using the QIAamp DNA Mini Kit (Qiagen). For whole exome sequencing, genomic DNA libraries were prepared using the SureSelectXT Reagent Kit (Agilent Technologies) and exome capture was performed using the SureSelectXT Human All Exon V4 kit (Agilent Technologies). Exome libraries were subjected to next generation sequencing using HiSeq2500 platform (Illumina) with 100 bp paired-end reads.

Exome-seq analysis

Paired-end sequencing reads in FASTQ format were generated using the Illumina software. The quality assessment of sequencing was performed with NGS QC Toolkit (v2.3.1) [30], and the high-quality reads were aligned to the human reference genome (hg19) using Burrows-Wheeler Aligner (BWA, v0.7.5a)[31]. Picard (v1.99) (http://picard.sourceforge.net) was used to remove PCR duplicates. The steps to improve the quality of the alignment, base quality recalibration, and local realignment around indels were performed using Genome Analysis Toolkit (GATK, v2.7-2)[32,33]. Calling variants and joint genotyping for all 19 samples were performed using HaplotypeCaller of GATK. The variants in Variant Call Format (VCF) were filtered by applying the following criteria: AN (Total number of alleles in called genotypes) < 38, DP (Approximate read depth) < 10, QD (Variant Confidence/Quality by Depth) < 5, GQ (Genotype confidence quality) < 20. The variants were annotated using In-house software with the National Center for Biotechnology Information (NCBI), Reference Sequence (RefSeq, downloaded at December 3rd, 2013),[34] and dbSNP (Build 137)[35] databases. The candidate selection step used only missense variants that are not listed in Common SNPs in the UCSC genome browser[36], a subset consisting of uniquely mapped variants that appear in at least 1% of the population[32,37].

CRISPR-Cas9 knock-in

pLX-sgPOLA1 was created by cloning *POLA1* guide RNA (5′-gTTAGTGATCTGCAATGCTAA -3′) into the pLX-sgRNA vector (Addgene 50662). For *POLA1* L764S knock-in, 1 million HCT-116 or HeLa cells were nucleofected (using 4D-Nucleofector, Lonza) with two plasmids, pSpCas9(BB)-2A-Puro (Addgene 48139) and pLX-sgPOLA1, and the single-stranded oligo 5′-ACCCCAAGCAAACACTGAATCCAACAGGAAATGCTTTTTCCCCCTTTCTAAGTTA AATTTACCATAATGTTCCCAGCGATGTTAGTGATCTGCAATGCTGATGGAAGAACA TTTAGCTCACACATGATCTGCAAAATGAACTTGGCATCTTTCCAGGTGTGTTCCAA CAGGTATAACAGTTGAGAAGATTCACTGTACAG -3′. Afterwards, cells were exposed to 5 μM CD437 for 7-13 days to select for cells with *POLA1* L764S knock-in. Clones were then expanded through sequential transfers to 24-well, 6-well and 10 cm plates. To validate knock-in, genomic DNA sequence flanking L764S were amplified from isolated clones with 5′ AGCATTGGGATCAGTGGTATG-3′ and 5′-TGTAAAACGACGGCCAGTTTCTCCCAACCAGTTCTTCCT-3′, and Sanger sequenced with M13 forward primer 5′- TGTAAAACGACGGCCAGT. For crystal violet staining, cells on plates were stained with 0.05% crystal violet, 1% formaldehyde in 1× PBS for 20 min at room temperature, followed by several rinses with deionized water to remove free stain.

BrdU incorporation assay

Parental and CD437 resistant HCT-116 cells were pretreated with CD437 for 2 hours. Afterwards, BrdU was added to a final concentration of 10 μM and cells were incubated for another 2 hours. Cells were then washed twice with PBS and collected by trypsinization. Genomic DNA was purified using the QIAamp DNA Mini Kit (Qiagen). 1 μg of purified genomic DNA was mixed with 100 μL of 0.4 N NaOH and was incubated for 30 min at room temperature and then placed on ice. The alkaline treated DNA was then neutralized

with the addition of 110 μL of 1M Tris-HCl, pH 6.8. 50 ng of denatured DNA was dot blotted on a nitrocellulose membrane and fixed to the membrane by UV crosslinking (UVA) with a Stratalinker. Western blot was performed by incubating the membrane with a 1:2000 dilution of mouse anti-BrdU monoclonal (B2531, Sigma) in Tris-Buffered Saline (with 0.05% Tween-20 and 1% nonfat milk) for 1 hour at room temperature, followed by an incubation with a dilution of 1:5000 goat anti mouse secondary antibody in Tris-Buffered Saline (with 0.05% Tween-20) for 1 hour at room temperature.

Expression and purification of POLA1 protein from insect cells

pFastBac1-POLA1 (containing the catalytic fragment of POLA1 corresponding to residues 335-1257 appended with an N-terminal His-tag) was described previously[19]. Plasmids containing L700S and L764S mutants were generated by site-directed mutagenesis using the following respective primer pairs: 5′AAATTTCAGCAAAGGAATCGATTCGTTGTAAAAGCT, 5′AGCTTTTACAACGAATCGATTCCTTTGCTGAAATTT and 5′AAATGTTCTTCCATCAGCATTGCAGATCACTA, 5′TAGTGATCTGCAATGCTGATGGAAGAACATTT. The production of baculovirus and protein expression in Sf9 cells were carried out following vendor's instructions. The purification of POLA1 protein was performed with affinity chromatography on Ni-NTA column (Qiagen) followed by size exclusion chromatography on Superdex 200 (GE Healthcare) equilibrated in 10 mM Tris-HCl, pH 7.8, 500 mM NaCl, 1% glycerol, and 1 mM DTT. Final protein samples were concentrated using a centricon filter (Millipore) and frozen as single-use aliquots in 10 mM Tris-HCl, pH 7.8, 50 mM NaCl, 1% glycerol, and 1 mM DTT.

In vitro primer extension assay

The primer extension substrate was generated by mixing 100 μL of a 5 μM fluorescein labeled 15 nt RNA oligo (5′-Fluorescein-rGrGrArArArGrGrArCrGrArArArArCrA-3′), 100 μL of a 7.5 μM 25-40 nt DNA oligo (5′-(A)nCCGGTGTTTCGTCCTTTCC-3′), and 100 μL of 10× reaction buffer (200mM Tris-HCl, pH7.8, 100 mM $MgCl_2$, 20 mM DTT, and 500 mM NaCl). The mixture was heated to 75°C and gradually cooled to room temperature. We confirmed annealing by native PAGE. Three-fold serial dilutions of CD437 were first prepared in DMSO, and then diluted 20 fold in H_2O. 1 μL of diluted CD437 was preincubated with 3.6 μL of POLA1 protein (containing 20 ng POLA1) at room temperature for 15 minutes. Afterwards, CD437 treated POLA1 protein was mixed with 3 μL of the primer extension substrate, 2 μL of 10 mg/ml BSA, and 0.4 μL of 2.5 mM dNTP. The primer extension reaction was incubated at room temperature for 15 minutes, and then terminated via addition of 10 μL of gel loading buffer II (Life Technologies). The reaction products were resolved on a 15% TBE UREA gel and visualized using a Typhoon scanner.

Cell cycle synchronization and caspase-3 western blot

HeLa cells were treated with 2 mM thymidine for 16 hours to arrest the cells at G1/S boundary. Thymidine was washed out to release cells into S phase. For caspase-3 western blots, anti-caspase-3 rabbit monoclonal antibody (obtained from Cell Signaling #9665) was used.

Measurement of binding affinity between CD437 and POLA1

Binding of POLA1 to CD437 was monitored by the changes in CD437 fluorescence anisotropy and CD437 fluorescence intensity enhancement. 100 nM of CD437 was mixed with the indicated concentrations of POLA1 or POLA1 mutant proteins, in a final buffer of 10 mM Tris-HCl, pH7.8, 50 mM NaCl, 1% glycerol, 1 mM DTT, 0.1% DMSO. Working CD437 stocks were prepared as 1 μM intermediate stocks in the above buffer (with 1% instead of 0.1% DMSO) and used within 2 hours of preparation. The CD437/POLA1 mixtures were incubated for 5 minutes at room temperature prior to analysis. Samples were analyzed in 3 mm by 3 mm cuvettes, in a PTI Quantamaster Spectrafluorometer, equipped with two emission monochrometers and detectors, and Glan-Thompson polarizers. Excitation and emission wavelengths were optimized for CD437 under our conditions. Excitation of CD437 was achieved at 315 nm (5 nm band-pass) and fluorescence emission was monitored at 395 nm (10 nm band-pass) with long-pass colored glass filters (Edmund Optics, WG-320) placed before the emission monochrometers to block scattered light. Cuvette temperature was regulated at 20 °C. Fluorescence emission intensity, in both parallel (I_{VV}) and perpendicular (I_{VH}) polarizations, was acquired for three minutes and averaged. These averaged intensities were corrected for background signal intensity and differences in monochrometer and detector efficiencies (G-factor correction), producing I_{VV}' and I_{VH}', then converted to anisotropy values $(I_{VV}' - I_{VH}')/(I_{VV}' + 2 * I_{VH}')$ or total intensity $(I_{VV}' + 2 * I_{VH}')$. Binding isotherms were fit to a single site receptor binding model, using Levenberg-Marquardt nonlinear least squares methods, with bound and free fluorescence anisotropy or intensities as fit parameters. Fit values for free fluorescence anisotropy or fluorescence intensity were similar to the values for CD437 in the absence of POLA1. For plotting fluorescence intensity enhancement, intensities were normalized such that the intensity of 100 nM free CD437 was equal to 1.0. Fluorescence enhancement binding isotherms were used to deduce K_{DS}. Error reported for K_{DS} are those deduced from local nonlinear fitting uncertainty.

Source data for exome sequencing data is available under NCBI SRA accession SRP068238.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Holmes WF, Soprano DR, Soprano KJ. Comparison of the mechanism of induction of apoptosis in ovarian carcinoma cells by the conformationally restricted synthetic retinoids CD437 and 4-HPR. J Cell Biochem. 2003; 89:262–278. DOI: 10.1002/jcb.10505 [PubMed: 12704790]

2. Hsu SL, Yin SC, Liu MC, Reichert U, Ho WL. Involvement of cyclin-dependent kinase activities in CD437-induced apoptosis. Exp Cell Res. 1999; 252:332–341. DOI: 10.1006/excr.1999.4625 [PubMed: 10527623]

3. Schadendorf D, et al. Treatment of melanoma cells with the synthetic retinoid CD437 induces apoptosis via activation of AP-1 in vitro, and causes growth inhibition in xenografts in vivo. The Journal of Cell Biology. 1996; 135:1889–1898. DOI: 10.1083/jcb.135.6.1889 [PubMed: 8991099]

4. Sun SY, Yue P, Shroot B, Hong WK, Lotan R. Induction of apoptosis in human non-small cell lung carcinoma cells by the novel synthetic retinoid CD437. J Cell Physiol. 1997; 173:279–284. DOI: 10.1002/SICI)1097-4652(199711)173:2<279: :AID-JCP36>3.0.CO;2-8 [PubMed: 9365537]

5. Rees MG, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. Nat Chem Biol. 2016; 12:109–116. DOI: 10.1038/nchembio.1986 [PubMed: 26656090]

6. Lotan R. Receptor-independent induction of apoptosis by synthetic retinoids. J Biol Regul Homeost Agents. 2003; 17:13–28. [PubMed: 12757019]

7. Pfahl M, Piedrafita FJ. Retinoid targets for apoptosis induction. Oncogene. 2003; 22:9058–9062. DOI: 10.1038/sj.onc.1207109 [PubMed: 14663484]

8. Rishi AK, et al. Post-transcriptional regulation of the DNA damage-inducible gadd45 gene in human breast carcinoma cells exposed to a novel retinoid CD437. Nucleic Acids Res. 1999; 27:3111–3119. [PubMed: 10454607]

9. Hail N Jr, Lotan R. Synthetic retinoid CD437 promotes rapid apoptosis in malignant human epidermal keratinocytes and G1 arrest in their normal counterparts. J Cell Physiol. 2001; 186:24–34. DOI: 10.1002/1097-4652(200101)186:1<24: :AID-JCP1006>3.0.CO;2-S [PubMed: 11147811]

10. Cincinelli R, et al. A novel atypical retinoid endowed with proapoptotic and antitumor activity. J Med Chem. 2003; 46:909–912. DOI: 10.1021/jm025593y [PubMed: 12620066]

11. Cincinelli R, et al. Synthesis and structure-activity relationships of a new series of retinoid-related biphenyl-4-ylacrylic acids endowed with antiproliferative and proapoptotic activity. J Med Chem. 2005; 48:4931–4946. DOI: 10.1021/jm049440h [PubMed: 16033272]

12. Martin B, et al. Selective synthetic ligands for human nuclear retinoic acid receptors. Skin Pharmacol. 1992; 5:57–65. [PubMed: 1315557]

13. Sun SY, et al. Identification of receptor-selective retinoids that are potent inhibitors of the growth of human head and neck squamous cell carcinoma cells. Clin Cancer Res. 2000; 6:1563–1573. [PubMed: 10778990]

14. Sun SY, et al. Dual mechanisms of action of the retinoid CD437: nuclear retinoic acid receptor-mediated suppression of squamous differentiation and receptor-independent induction of apoptosis in UMSCC22B human head and neck squamous cell carcinoma cells. Mol Pharmacol. 2000; 58:508–514. [PubMed: 10953043]

15. Hsu CA, et al. Retinoid induced apoptosis in leukemia cells through a retinoic acid nuclear receptor-independent pathway. Blood. 1997; 89:4470–4479. [PubMed: 9192771]

16. Parrella E, et al. Antitumor activity of the retinoid-related molecules (E)-3-(4′-hydroxy-3′-adamantylbiphenyl-4-yl)acrylic acid (ST1926) and 6-[3-(1-adamantyl)-4-hydroxyphenyl]-2-naphthalene carboxylic acid (CD437) in F9 teratocarcinoma: Role of retinoic acid receptor gamma and retinoid-independent pathways. Mol Pharmacol. 2006; 70:909–924. DOI: 10.1124/mol. 106.023614 [PubMed: 16788091]

17. Wacker SA, Houghtaling BR, Elemento O, Kapoor TM. Using transcriptome sequencing to identify mechanisms of drug action and resistance. Nat Chem Biol. 2012; 8:235–237. DOI: 10.1038/nchembio.779 [PubMed: 22327403]

18. Glaab WE, Tindall KR. Mutation rate at the hprt locus in human cancer cell lines with specific mismatch repair-gene defects. Carcinogenesis. 1997; 18:1–8. [PubMed: 9054582]

19. Baranovskiy AG, et al. Structural basis for inhibition of DNA replication by aphidicolin. Nucleic Acids Res. 2014; 42:14013–14021. DOI: 10.1093/nar/gku1209 [PubMed: 25429975]

20. Mishur RJ, Griffin ME, Battle CH, Shan B, Jayawickramarajah J. Molecular recognition and enhancement of aqueous solubility and bioactivity of CD437 by beta-cyclodextrin. Bioorganic & medicinal chemistry letters. 2011; 21:857–860. DOI: 10.1016/j.bmcl.2010.11.073 [PubMed: 21185186]

21. Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. Mol Cell. 2010; 40:179–204. DOI: 10.1016/j.molcel.2010.09.019 [PubMed: 20965415]

22. Valli C, et al. Atypical retinoids ST1926 and CD437 are S-phase-specific agents causing DNA double-strand breaks: significance for the cytotoxic and antiproliferative activity. Molecular cancer therapeutics. 2008; 7:2941–2954. DOI: 10.1158/1535-7163.mct-08-0419 [PubMed: 18790775]

23. Finn RS, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. The Lancet Oncology. 2015; 16:25–35. DOI: 10.1016/s1470-2045(14)71159-3 [PubMed: 25524798]

24. Ikegami S, et al. Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase-alpha. Nature. 1978; 275:458–460. [PubMed: 692726]

25. Sessa C, et al. Phase I and clinical pharmacological evaluation of aphidicolin glycinate. J Natl Cancer Inst. 1991; 83:1160–1164. [PubMed: 1886148]

26. Edelson RE, Gorycki PD, MacDonald TL. The mechanism of aphidicolin bioinactivation by rat liver in vitro systems. Xenobiotica. 1990; 20:273–287. DOI: 10.3109/00498259009046847 [PubMed: 2110702]

27. Prasad G, Edelson RA, Gorycki PD, Macdonald TL. Structure-activity relationships for the inhibition of DNA polymerase alpha by aphidicolin derivatives. Nucleic Acids Res. 1989; 17:6339–6348. [PubMed: 2505232]

28. Sun SY, et al. Mechanisms of apoptosis induced by the synthetic retinoid CD437 in human non-small cell lung carcinoma cells. Oncogene. 1999; 18:2357–2365. DOI: 10.1038/sj.onc.1202543 [PubMed: 10327056]

29. Pedrali-Noy G, et al. Synchronization of HeLa cell cultures by inhibition of DNA polymerase alpha with aphidicolin. Nucleic Acids Res. 1980; 8:377–387. [PubMed: 6775308]

30. Patel RK, Jain M. NGS QC Toolkit : a toolkit for quality control of next generation sequencing data. PLoS One. 2012; 7:e30619. [PubMed: 22312429]

31. Hashiguchi K, et al. Involvement of ETS1 in thioredoxin-binding protein 2 transcription induced by a synthetic retinoid CD437 in human osteosarcoma cells. Biochem Biophys Res Commun. 2010; 391:621–626. DOI: 10.1016/j.bbrc.2009.11.109 [PubMed: 19932085]

32. Danecek P, et al. The variant call format and VCFtools. Bioinformatics. 2011; 27:2156–2158. DOI: 10.1093/bioinformatics/btr330 [PubMed: 21653522]

33. McKenna A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297–1303. DOI: 10.1101/gr.107524.110 [PubMed: 20644199]

34. Pruitt KD, et al. RefSeq: an update on mammalian reference sequences. Nucleic Acids Res. 2014; 42:D756–763. DOI: 10.1093/nar/gkt1114 [PubMed: 24259432]

35. Sherry ST, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001; 29:308–311. [PubMed: 11125122]

36. Karolchik D, et al. The UCSC Genome Browser database: 2014 update. Nucleic Acids Res. 2014; 42:D764–770. DOI: 10.1093/nar/gkt1168 [PubMed: 24270787]

37. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. DOI: 10.1086/519795 [PubMed: 17701901]
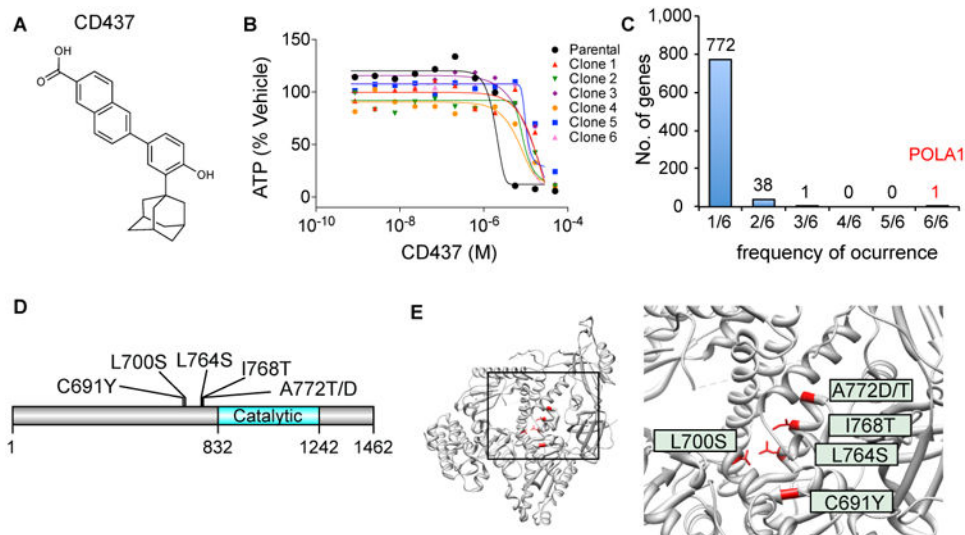
**Figure 1. Mutations in *POLA1* render resistance to CD437**

(A) Chemical structure of CD437.

(B) CD437 dose responses of the proliferation of parental and six independent CD437 resistant clones (72 hours of treatment). Each point represents the average of two biological replicates. IC_{50}s are 3 μM (parental), 32 μM (clone 1), 15 μM (clone 2), 13 μM (clone 3), 11 μM (clone 4), 17 μM (clone 5), and 18 μM (clone 6).

(C) *POLA1* is mutated in all six exome-sequenced clones.

(D) CD437 resistant mutations are clustered in POLA1, outside of its catalytic center.

(E) CD437 resistant mutations highlighted in red on the crystal structure of human POLA1 (PDB: 4QCL).
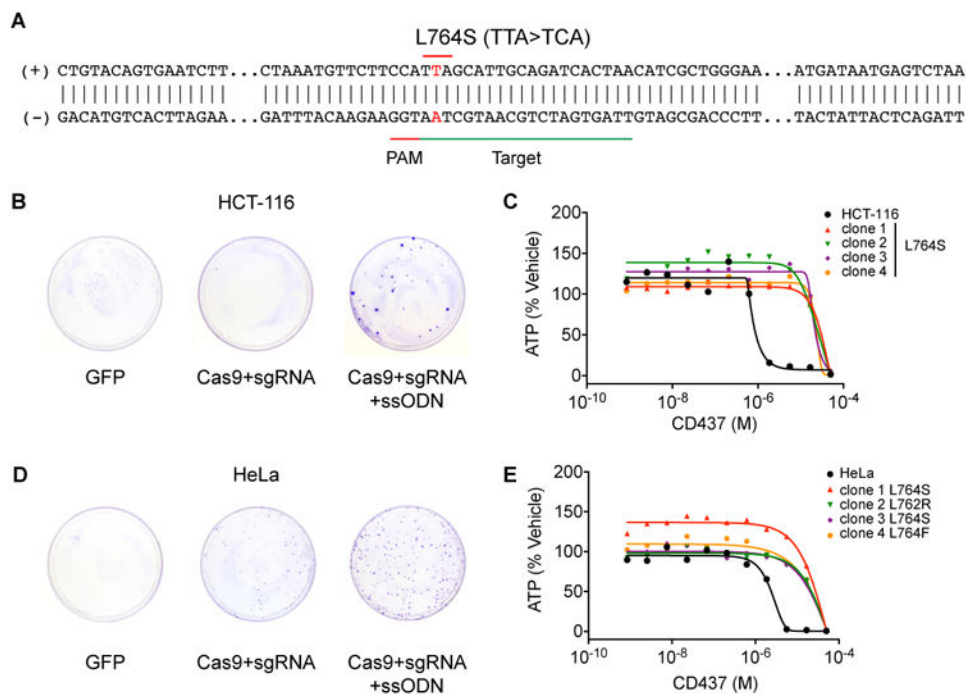
Figure 2. *POLA1* L764S knock-in using CRISPR/Cas9 technology results in CD437 resistance

(A) The edited genomic sequence of *POLA1*, highlighting the PAM motif, the target sequence, and L764S mutation.

(B) *POLA1* L764S genome-editing in HCT-116 cells gives rise to CD437 resistance visualized by crystal violet staining. Mock transfection with either GFP plasmid or omission of the repair template results in no CD437 resistance.

(C) CD437 dose responses of parental HCT-116 and isolated *POLA1* L764S knock-in clones (72 hours of treatment). Each point represents the average of two biological replicates. IC_{50}s are 0.8 μM (parental), 34 μM (clone 1), 18 μM (clone 2), 21 μM (clone 3) and 23 μM (clone 4).

(D) Similar experiment as (B) performed with HeLa cells.

(E) CD437 dose responses of parental HeLa cells and isolated *POLA1* L764S knock-in clones (72 hours of treatment). Each point represents the average of two biological replicates. IC_{50}s are 2.5 μM (parental), 34 μM (clone 1), 38 μM (clone 2), 21 μM (clone 3) and 20 μM (clone 4).
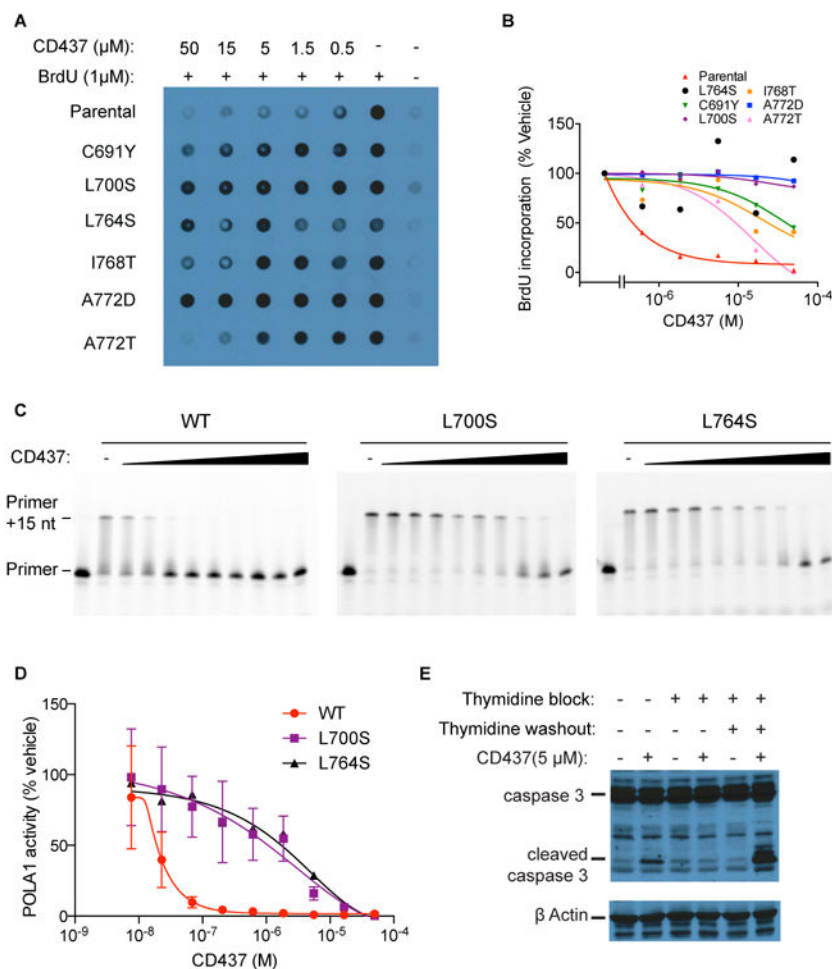
Figure 3. CD437 inhibits DNA replication *in vivo* and POLA1 activity *in vitro*

(A) CD437 inhibits BrdU incorporation in parental cells visualized by dot blot of genomic DNA. Cells expressing indicated *POLA1* mutations are resistant to CD437. These experiments were performed twice and a representative result is shown.

(B) Quantification of (A).

(C) CD437 inhibits POLA1 primer extension activity *in vitro*. L700S and L764S mutants of POLA1 are less sensitive to CD437. Primer extension was determined in the presence of increasing concentrations of CD437 (0.0076 μM, 0.023 μM, 0.069 μM, 0.21 μM, 0.62 μM, 1.9 μM, 5.6 μM, 17 μM, 50 μM). The "-" represents a DMSO control. The activity of POLA1 is determined by the detection of primer extension of a 30 nucleotide product (15 dNTP incorporations). This experiment was conducted three times and a representative gel is shown (two other replicates are shown in Supplementary Figure 4).

(D) Quantification of intensities of 30 nucleotide primer extension product for each lane in (C) and Supplementary Figure 4. The error bars reflect SEM. IC$_{50}$s for the wild type, L700S, and L764S POLA1 are 21±1.5 nM, 1.5±1.2 μM, and 3.1±1.9 μM, respectively.

(E) CD437 is a S-phase toxin in HeLa cells. HeLa cells were arrested at G1 phase by the addition of 2 mM thymidine, and released into S phase by thymidine washout. Asynchronous, G1 arrested cells, and S phase cells were treated with 5 μM CD437 or

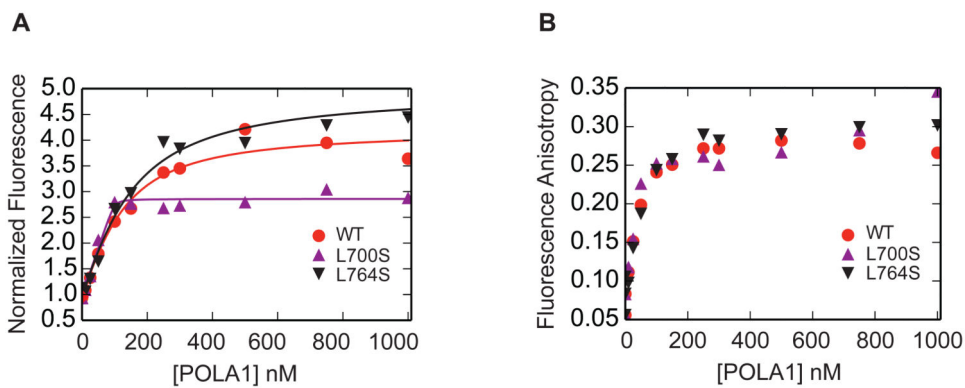DMSO for 8 hours. Lysates from these cells were analyzed by immunoblotting for caspase-3. Full gels are displayed in Supplementary Figure 6.

Figure 4. CD437 binds POLA1 *in vitro*

(A) POLA1 enhances the fluorescence intensity of CD437. Binding isotherms were fit to a single site receptor binding model with bound and free fluorescence intensities as fit parameters. Dissociation constants (Kd) of WT, L700S, and L764S are 67 ± 22 nM, <5 nM, 94 ± 18 nM, respectively. This experiment was conducted twice and a representative result is shown (a technical replicate is shown in Supplementary Figure 7A.

(B) POLA1 enhances the fluorescence anisotropy of CD437. This experiment was conducted twice and a representative result is shown (a technical replicate is shown in Supplementary Figure 7B).